DOCUMENT RESUME

ED 411 309 TM 027 360

AUTHOR Katz, Sandra N.; Hall, Ellen; Lesgold, Alan
TITLE Cognitive Task Analysis and Intelligent Computer-Based Training Systems: Lessons Learned from Coached Practice Environments in Air Force Avionics.
INSTITUTION Pittsburgh Univ., PA. Learning Research and Development Center.; Armstrong Lab, Brooks AFB, TX. Human Resources Directorate.
PUB DATE 1997-03-00
NOTE 23p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).
PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Aviation Mechanics; Cognitive Processes; Cooperation; *Intelligent Tutoring Systems; Job Analysis; *Job Skills; *Systems Development; Task Analysis; *Test Construction

ABSTRACT

This paper describes some results of a collaborative effort between the University of Pittsburgh and the Air Force to develop advanced troubleshooting training for F-15 maintenance technicians. The focus is on the cognitive task methodology used in the development of three intelligent tutoring systems to inform their instructional content and approach, and how task analysis results are reflected in particular features of one of these tutors. A well-conducted cognitive task analysis (CTA) can give a system developer information about the knowledge and skills students find difficult. In the three developed systems, a CTA methodology called PARI (Precursor (goal), Action, Result, and Interpretation) (3) informed the design of coaching and postproblem reflection. These systems, Sherlock, Hydrive, and Eaglekeeper, are designed to give those who maintain F-15 aircraft feedback about their reasoning errors and violations of good troubleshooting practice. Examples show how the PARI methodology informed the development of these systems, and results with 18 tutored novices, 23 untutored novices, and 13 master technicians are presented to show the efficacy of the training. The two fully developed systems are proving efficient and practical in improving student performance. The third, Eaglekeeper, remains in an earlier stage of development. Work with these systems illustrates the importance of early CTA to save time and effort in system development. (Contains 10 figures.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 411 309

Cognitive Task Analysis and Computer-based Training Systems: Lessons Learned from Coached Practice Environments in Air Force Avionics

Sandra N. Katz
University of Pittsburgh
Learning Research and Development Center
Ellen M. Hall
Air Force Armstrong Laboratory

This paper describes some results of a collaborative effort between the University of Pittsburgh and the Air Force to develop advanced troubleshooting training for F-15 maintenance technicians. The focus of this presentation is on the cognitive task analysis methodology that was used in the development of three intelligent tutoring systems to inform their instructional content and approach, and how task analysis results are reflected in particular features of one of these tutors.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sandra Katz

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

TM027360

BEST COPY AVAILABLE

Cognitive Task Analysis and Intelligent Computer-based Training Systems: Lessons Learned from Coached Practice Environments in Air Force Avionics

Sandra Katz, Ellen Hall, and Alan Lesgold

Coached practice environments immerse students in complex tasks similar to those they might face on the job. They simulate the job environment and support trainees in solving problems somewhat beyond their ability. Developers of coached practice environments need to know what instruction should focus on, out of the vast sea of knowledge and skills that could be taught. This requires an understanding of expertise in the task domain: i.e., what knowledge and skills do experts draw upon when they face unfamiliar, challenging problems? Equally important is an understanding of which concepts and skills students typically have difficulty acquiring, and what types of tasks are difficult for them.

A well-conducted cognitive task analysis (CTA) can provide this information to the system developer, and the CTA can be critical to the effectiveness of the tutoring system. For example, research by one of the authors showed that a coached practice environment for aircraft hydraulics maintenance which was informed by a CTA showed a significant learning effect, particularly for problems requiring students to develop their own troubleshooting strategies rather than follow set procedures (1). By contrast, a tutor for the same domain which lacked the benefit of CTA showed no learning effect. The critical difference between the systems rested in what they focused on. The CTA-informed tutor focused on strategy and the reasons for carrying out actions. It did not target procedural skills because the CTA showed no expert-novice differences in procedural skill. The CTA-deprived tutor, in contrast, focused mainly on procedural skill.

In this paper, we discuss the types of information that a CTA can provide to developers of coached practice environments, and illustrate how this information shaped the implementation of training systems for Air Force aircraft maintenance—namely, *Sherlock*, *Hydrive*, and *EagleKeeper*. These systems were developed by the authors (2). Our discussion focuses on how a CTA methodology called PARI (3) informed the design of coaching and post-problem reflection ("debrief") in these systems. We also discuss the limitations of CTA for tutor design. In particular, the PARI-based CTA does not address presentation issues, such as how to provide advice or give feedback. Nor does the CTA reveal what the criteria of expert performance are. We discuss how other information-gathering methods can fill these gaps—e.g., continuous interaction with subject matter experts, policy-capturing analyses, and observational studies of students using the system with assistance from a human tutor—and demonstrate how the information acquired through these techniques was incorporated in our systems.

(1) Hall, E.P., Rowe, A.L., Pokorny, R.A., & Boyer, B.S. (1996). A field evaluation of two intelligent tutoring systems. Paper presented at the annual meeting of the American Educational Research Association, April 9 - 12, New York, NY.

(2) See for example: Gott, S.P., Lesgold, A., & Kane, R.S. (1996). Tutoring for transfer of technical competence. In B.G. Wilson, Constructivist Learning Environments (pp. 33-48). Englewood Cliffs, NJ: Educational Technology Publications.

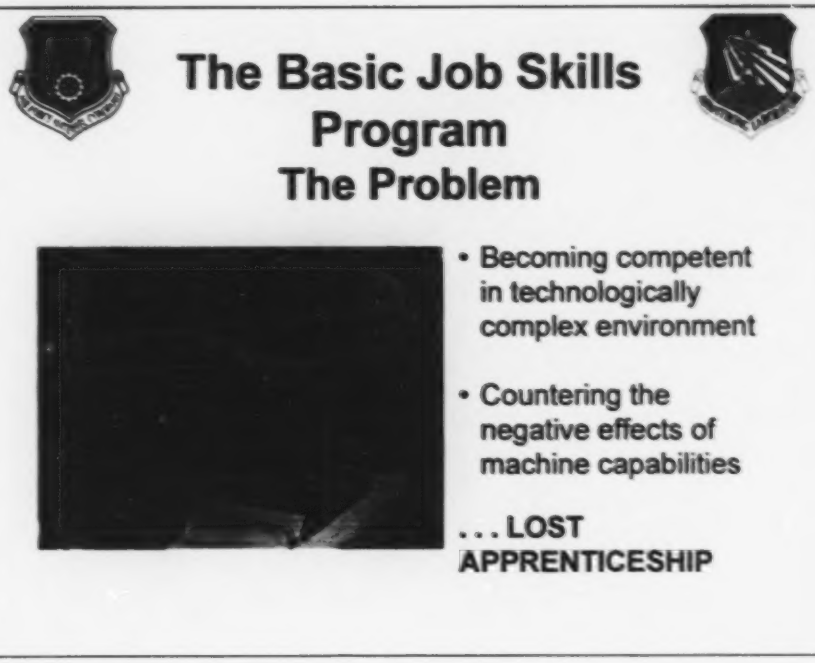
(3) PARI stands for Precursor (goal), Action, Result, and Interpretation. This is the information that experts are prompted for during the CTA as they work through a problem. Hall, E.P., Gott, S.P., & Pokorny, R.A. (1995). A procedural guide to cognitive task analysis: The PARI methodology. AL/HR TR-1995-0108. Brooks AFB TX: Human Resources Directorate, Manpower and Personnel Research Division.

TM027360

Overview

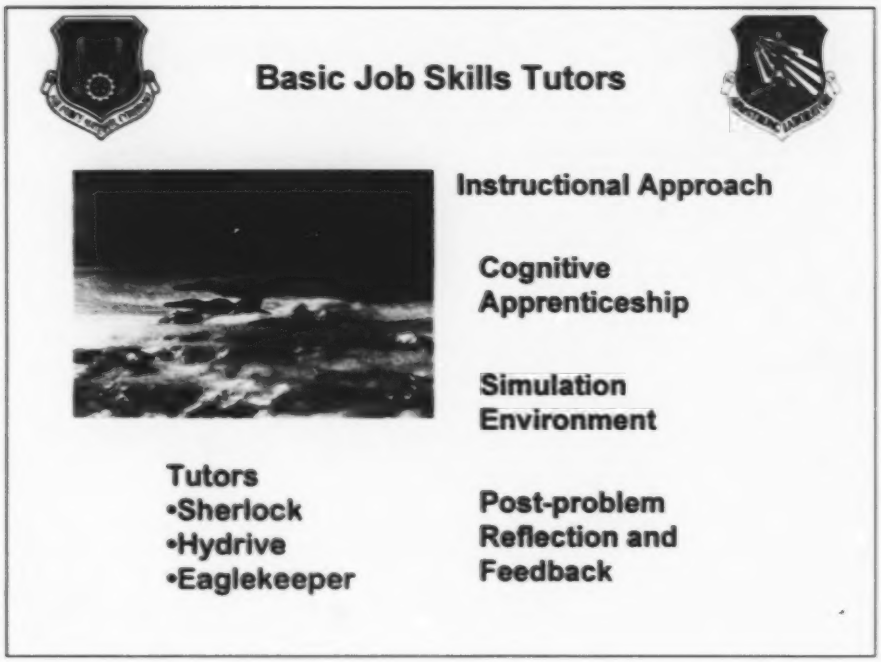
- The Basic Job Skills Program
- Cognitive Task Analysis Approach
- How Task Analysis Informs Tutor Development
- Lessons Learned

I'm Ellen Hall from the Armstrong Laboratory and I'll begin the presentation with a brief description of the Basic Job Skills Program under which this work was conducted. I'll then describe the task analysis methodology that was referred to earlier, and show some evidence that the use of this methodology is related to enhanced training effectiveness in the tutors. Then Sandra Katz from LRDC will talk about some specific examples of how the task analysis data were used in developing the instruction and some of the lessons we learned about the process of tutor development based on task analysis data.

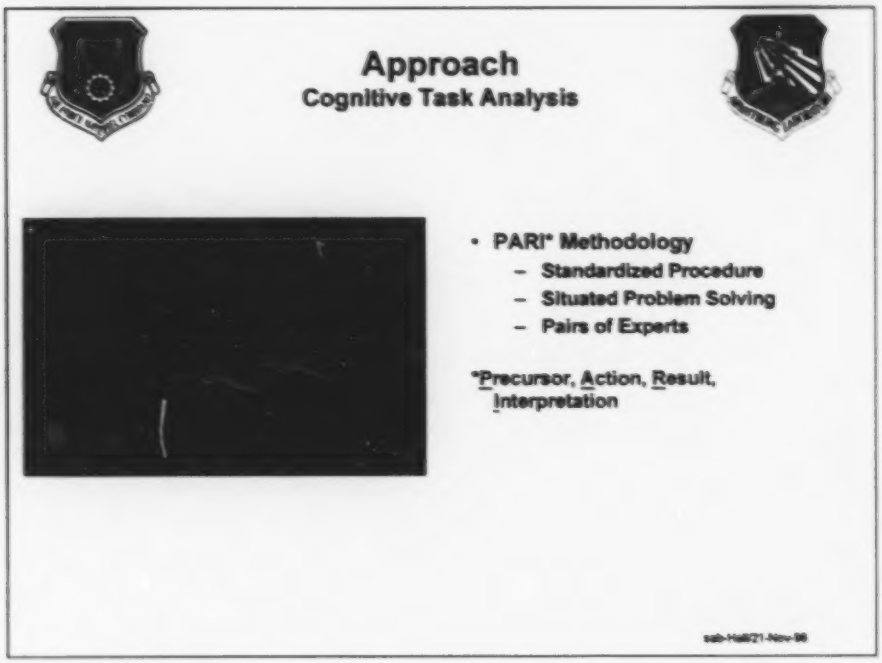

The Basic Job Skills Program
The Problem
• Becoming competent in technologically complex environment
• Countering the negative effects of machine capabilities
. . . LOST APPRENTICESHIP

The Basic Job Skills Program was initiated to address a problem that we're seeing more and more frequently as technology is introduced into the workplace to make the technician's job "easier." In the maintenance world, for example, software diagnostics enable technicians to isolate faults in complex systems without necessarily having to understand the fix and how the diagnosis was made. While such job aids are highly effective in many cases, they are not 100% reliable in the sense that they cannot diagnose every conceivable equipment failure. That degree of reliability would require the developers of the diagnostics to anticipate everything that could possibly go wrong with the system which is virtually impossible for some systems given their complexity. So what we end up with are technicians who have come to rely on these aids to get their job done (since they work most of the time), and who, as a result, have lost the learning opportunities associated with troubleshooting those faults on their own. They are thus ill-equipped to solve the most difficult problems that arise when these aids fail.

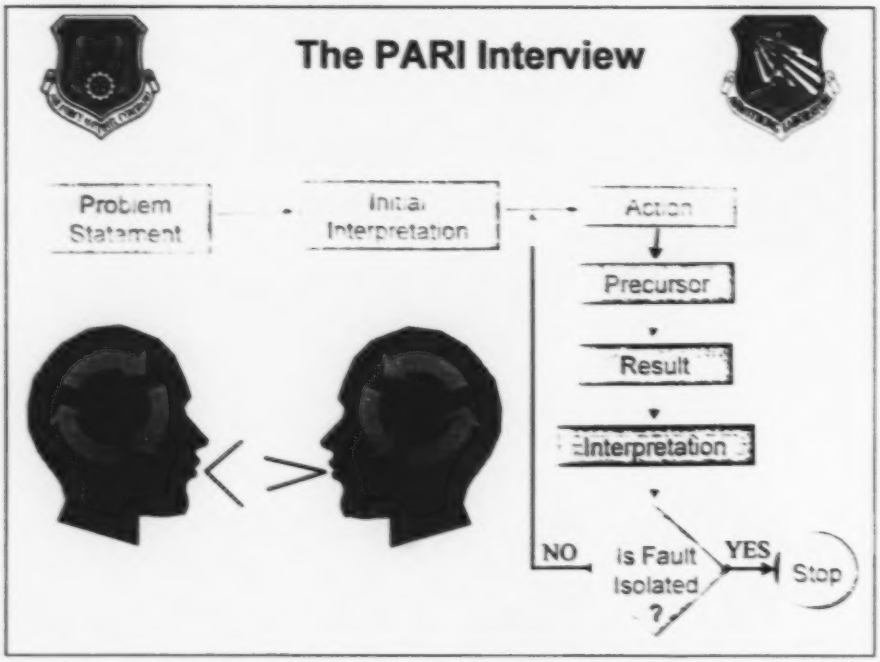
BEST COPY AVAILABLE



Intelligent tutoring systems were seen as a means of restoring these learning opportunities by providing coached practice in troubleshooting faults on a simulation of the equipment. The tutors also provide students with the opportunity to reflect on their troubleshooting performance during post-problem reflection where the tutor provides specific feedback about reasoning errors and violations of good troubleshooting practice. The three tutors listed on this slide are at various stages of transition to the F-15 maintenance community. Sherlock was the first tutor to be transitioned back in '94. It was developed by LRDC and targets the manual avionics test station specialty. Hydrive was transitioned in '95; it targets F-15 hydraulics troubleshooting and was developed by Educational Testing Service. Both Sherlock and Hydrive were evaluated in controlled field tests and I'll be talking about the results of those evalutions in just a moment. Eaglekeeper is currently under development by LRDC and targets flightline avionics troubleshooting.

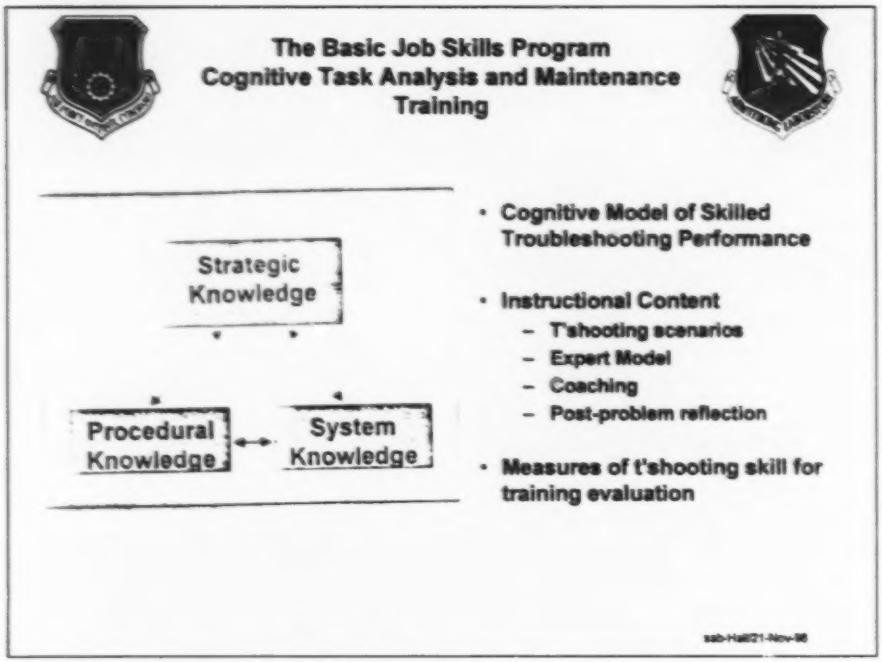


The task analysis methodology that informed these tutors was designed to elicit the knowledge and reasoning that underlies skilled troubleshooting. The PARI methodology is a standardized method for conducting interviews in a structured way to elicit these skills. PARI is an acronym that stands for the four elements of this structure, which I'll describe in just a moment. The standardized structure streamlines the interview process and makes it possible to easily compare the results of interviews with different experts and make comparisons between expert and novice troubleshooting performance. The second feature of the interview method is that the interviews are situated in the context of solving actual troubleshooting problems. The idea here is that it's easier for experts to articulate that knowledge when it's being activated to solve a problem. So experts aren't just telling us what they need to know to do they're jobs, they're showing us how they're using that knowledge to solve the problem. The third feature of the method was suggested by Allen Collins when he was consulting on the project during the early stages of development of this procedure. During the interview, pairs of experts interact during a verbal simulation of a troubleshooting scenario, with one expert acting as the problem solver, and the second expert essentially simulating the equipment that the problem solver is interacting with. So the second expert knows the location fault and can tell the first expert how the equipment will respond at each step of the troubleshooting solution as the problem solver takes measurements, or replaces components, and interacts with the equipment.



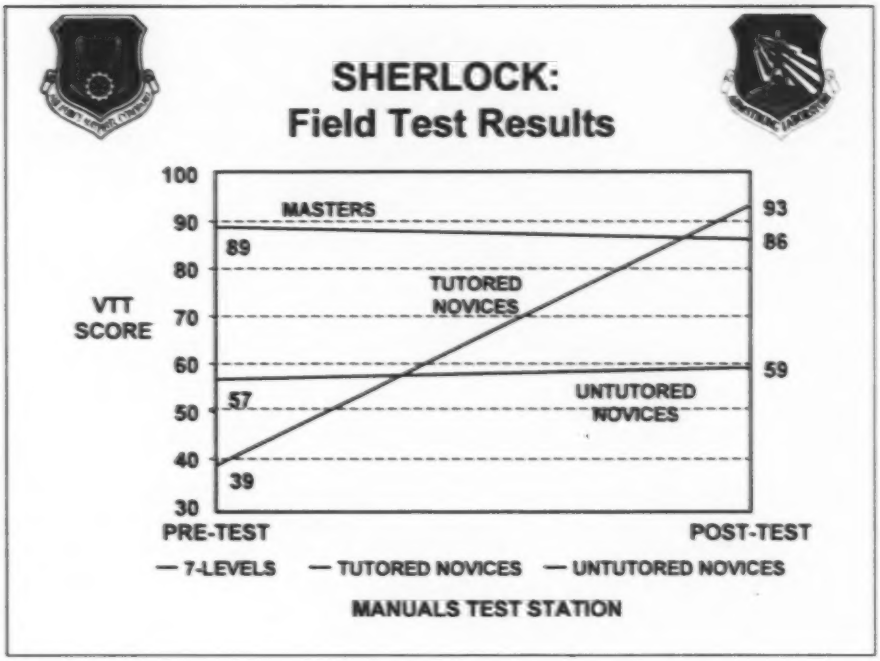
This slide shows the PARI interview structure. To begin the interview, the problem solver is presented with a problem statement which describes a set of symptoms indicating a fault in a piece of equipment. S/he is then required to specify, step by step, the actions s/he would take to solve the problem. At each step of the solution, four pieces of information are elicited which correspond to the elements of the PARI structure: the first piece of information is the cognitive **P**recursor to the action, or the goal of the action at that step; the second piece is the **A**ction itself; the third piece is the **R**esult of the action at that step in terms of the equipment response, and that information is provided by expert number two. The fourth piece is the expert's **I**nterpretation of the result in terms of the precursor or goal at that step. This probe structure is repeated at each troubleshooting step and the interview continues until the fault is isolated. Then several reviews of the problem solution are conducted to elaborate in various ways on the elements of each step. For example, in one of these reviews the problem solver is asked to name alternative actions he could have chosen at each step to pursue the stated goal, and then to contrast those actions with the action chosen in terms of the costs and benefits of each. The idea of these reviews is to elicit the decision factors that influence the selection of troubleshooting actions and goals, and to capture the mental models that underly the interpretation of results.

BEST COPY AVAILABLE



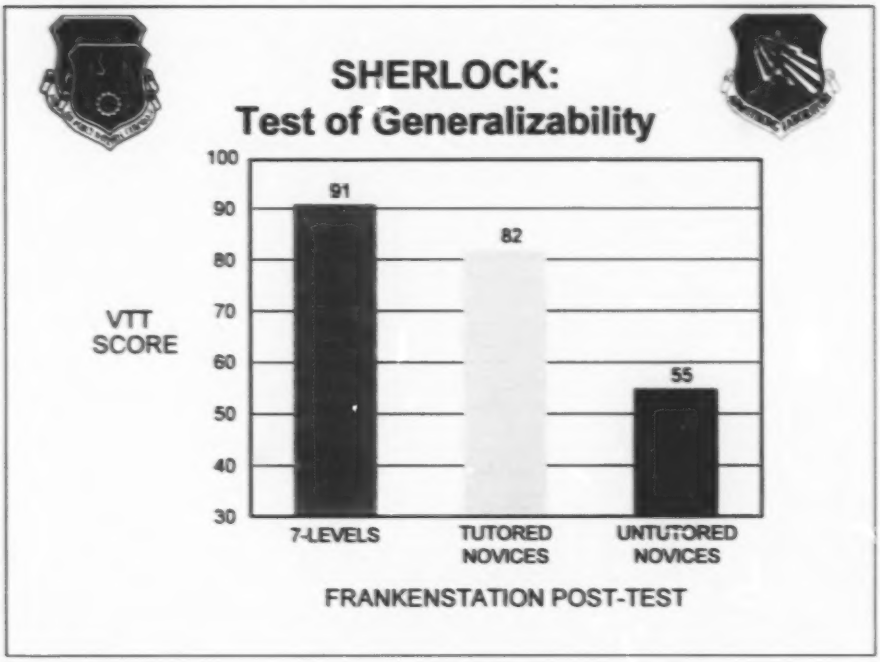
The troubleshooting protocols that result from presenting these scenarios to technicians ranging from experts to novices inform the tutors at a number of levels. On the left side of this slide you see in its most abstract form the cognitive model of skilled troubleshooting that informs all of these components of the tutor (listed on the right side of the slide under "Instructional Content"). This model represents the three types of knowledge that are coordinated during skilled troubleshooting and are captured in the PARI interviews. Procedural knowledge is knowledge of how to carry out troubleshooting actions such as taking measurements, repairing cables, or swapping out components. It's the easiest type of knowledge for technicians to develop because it's associated with observable behaviors. The second type of knowledge to develop is understanding how the system works and it's acquired at first by exercising the procedural knowledge to interact with the equipment and observe its behavior. The third and last type of knowledge to develop is strategic knowledge and it has been defined as knowledge of what to do and when to do it. It serves an executive control function and is very much dependent on having the system and procedural knowledge available to make those decisions. As a result, it appears to develop after last after many years of experience. This was the general model that informed our tutor development efforts, and now I'd like to show you the results of our field tests of Sherlock and Hydrive. Following that, Sandy will show you some specific examples of how the PARI data informed the development of Sherlock.

BEST COPY AVAILABLE



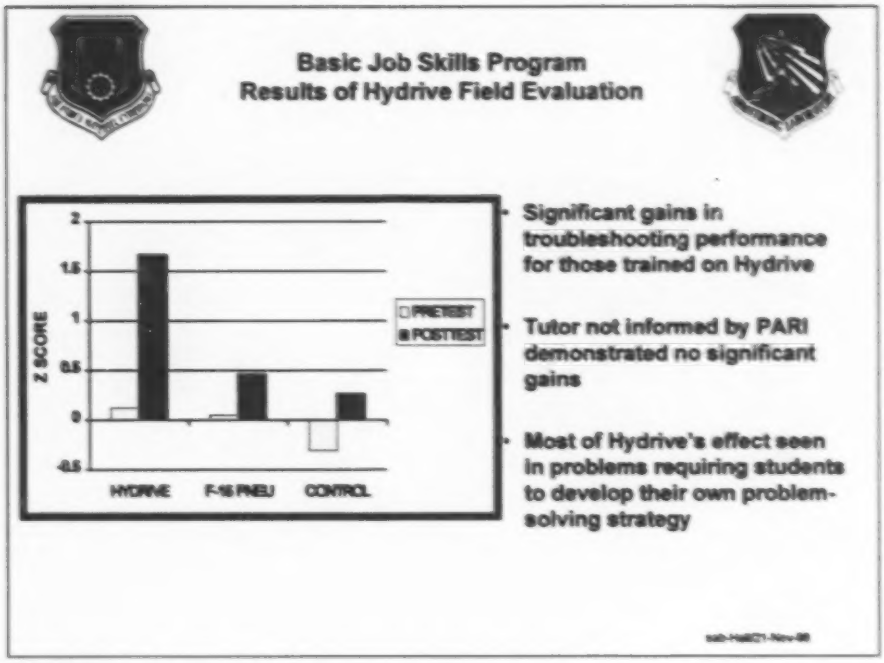
This slide shows pre- and post-test results for three groups of F-15 manual avionics test station technicians who participated in the field evaluation of Sherlock. The tutored novices (n=18) and untutored novices (n=23) showed no statistically significant differences prior to the intervention on measures of troubleshooting proficiency (VTT, or verbal troubleshooting test score), aptitude (ASVAB electronics composite score), or experience. The master technicians (n=13) had over four times the job-related experience as the novices and significantly higher troubleshooting proficiency scores prior to the tutoring phase of the study. During the tutoring phase, tutored novices received an average of 20 hours of training on Sherlock over a period of 3 weeks while the other two groups continued their normal duty assignments. The post-test verbal troubleshooting scores were significantly higher for the tutored novices compared to the untutored novices (VTT_3, $t[39]=-4.04$, $p<.001$; VTT_4, $t[39]=$

-3.72, $p<.001$) and were comparable to those of the master technicians.



In order to determine whether the troubleshooting skills acquired through tutoring on Sherlock were the type of flexible skills needed to deal with completely novel troublehshooting situations, a test of generalizability was constructed that required technicians to verbally isolate faults on a test station they had no familiarity with. This test station was conceived in the laboratory by one of our subject-matter experts and for that reason was called "frankenstation." Although similar in function to the manual test station these technicians used on their jobs, frankenstation was a computer-controlled test station, so technicicians had to troubleshoot it by routing signals electronically rather than manually, so the procedural knowledge required was very different from their own job. The question was whether technicians could transfer the strategic and system knowledge acquired in Sherlock to this novel troubleshooting environment. This slide shows the mean verbal troubleshooting scores of the three groups of technicians (tutored novices, n=17; untutored novices, n=21; master technicians, n=12) on the frankenstation problems. Again the tutored novices significantly outperformed the untutored novices ($t[36]$ -2.93, $p<.01$)and their scores were comparable to those of the master technicians.

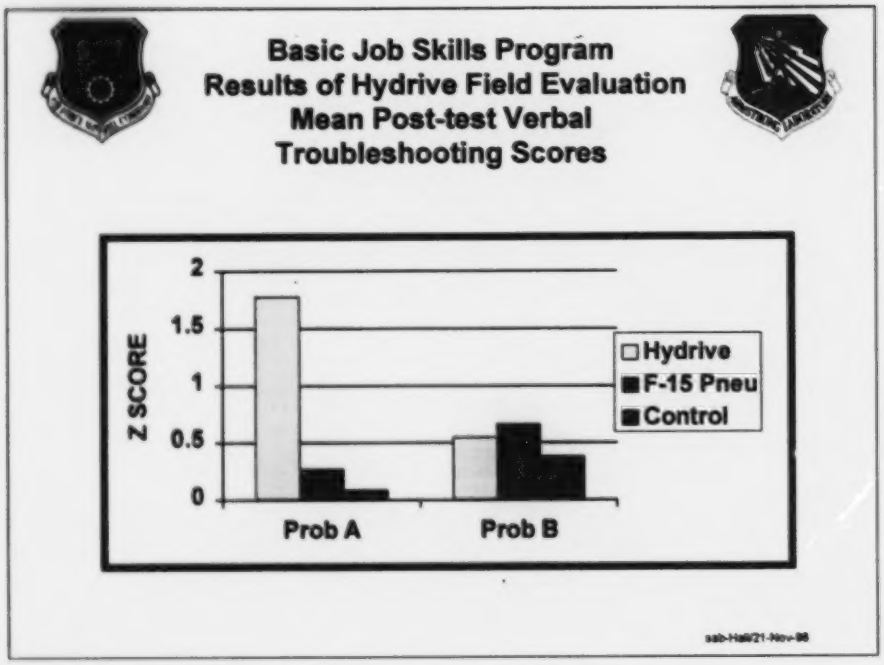
Overall, the Sherlock results show that the the tutor was in fact effective in enhancing troubleshooting proficiency, and that students acquired skills that went beyond the those based on knowledge of observable procedures; the skills that generalized to the frankenstation task were those based on system and strategic knowledge.



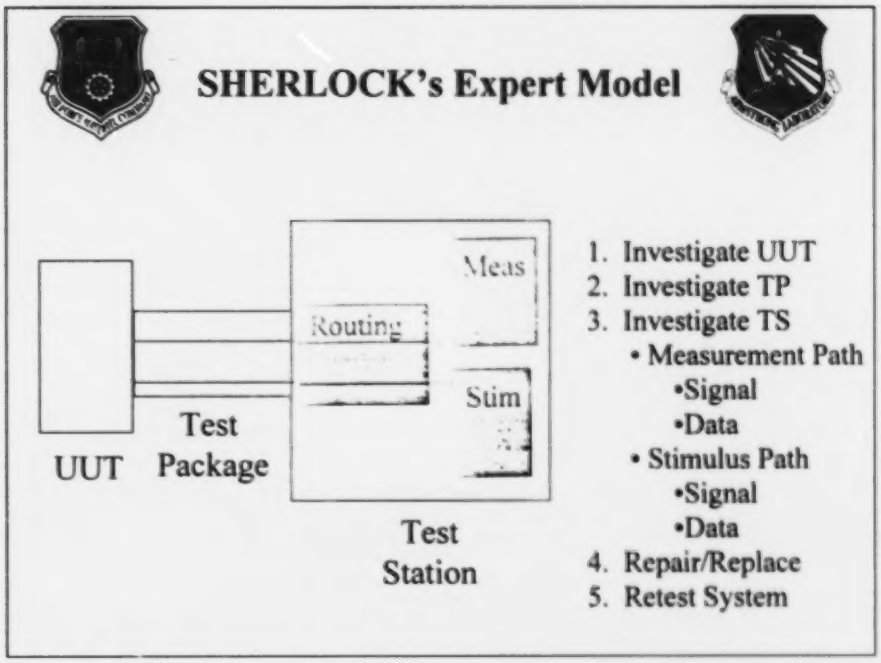
In the field evaluation of Hydrive, we had the opportunity to compare the effects of two different intelligent tutoring systems on the troubleshooting performance of F-15 hydraulics technicians. One critical difference between Hydrive and F-15 Pneudraulics tutor was that Hydrive was informed by a PARI analysis, while the F-15 Pneudraulics tutor relied on input from a single subject-matter expert in development of the instructional content. While both tutors contained the same set of troubleshooting scenarios (at least for the purpose of the field test), Hydrive's instruction focused primarily on the strategic and system knowledge underlying expert troubleshooting in this domain. Procedural knowledge was much less emphasized (however, safety procedures were emphasized) because the task analysis demonstrated that it was not procedural knowledge that distinguished expert and novice technicians. While both groups demonstrated knowledge of troubleshooting procedures, only experts demonstrated the strategic knowledge that led to efficient and effective troubleshooting. Instruction in the F-15 Pneudraulics tutor, on the other hand, focused mainly on procedures, and to a smaller extent, system knowledge.

This slide shows the pre- and post-test results on a verbal troubleshooting test that compared novice technicians tutored on Hydrive or the F-15 Pneudraulics Tutor, and a third control group who continued with their normal job duties during the tutoring phase of the study. While Hydrive students improved significantly from pre- to post-test ($t[19]=4.14$, $p<.001$), those in the other two groups showed no significant improvement.

BEST COPY AVAILABLE

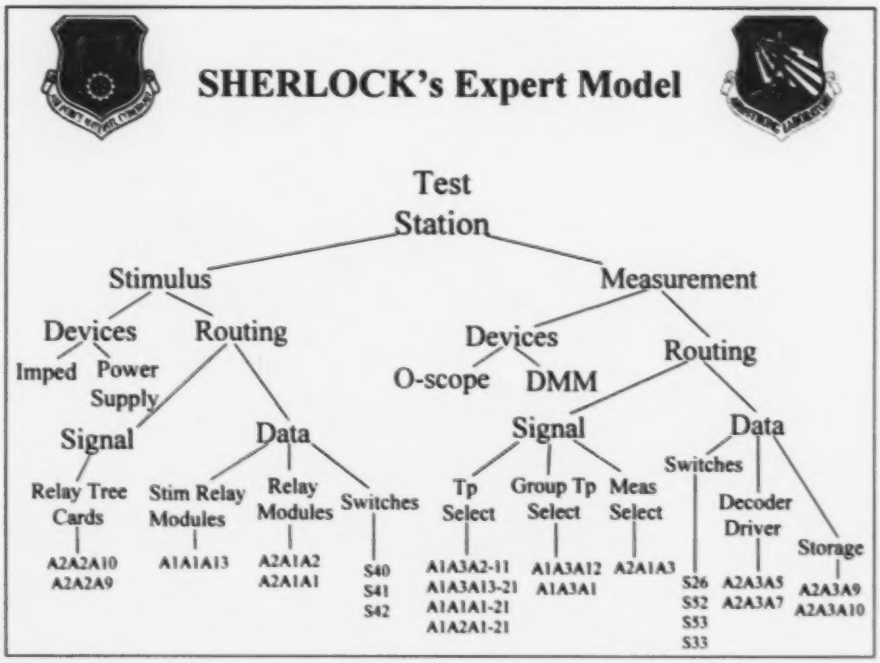


Further, when post-test performance was analyzed by the type of problem being solved we found that most of Hydrive's effect was seen on problems requiring technicians to develop their own troubleshooting strategies (Problem A). On Problem B, fault isolation guides were available that would have led to the solution of these problems, and once students had that, no group had any particular advantage over another. Thus, the model of skilled troubleshooting that informed these tutors provided a useful framework to guide the instruction and target those skills that distinguish experts from novices.

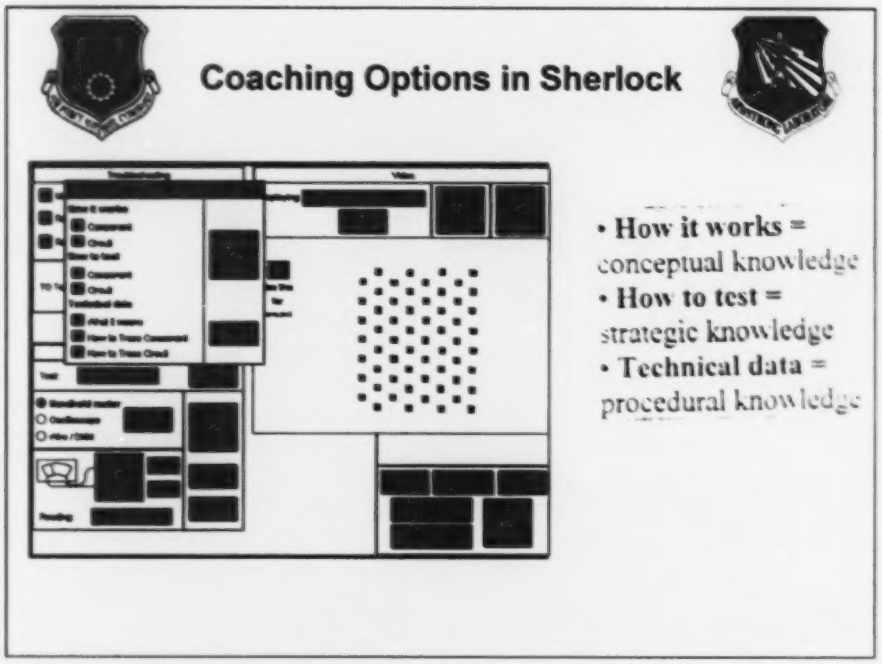


The PARI interview data enabled us to derive a model of the basic goals that experts achieve on the route towards isolating the faulty component: Investigate the UUT, Investigate the TP, Investigate the TS (the Measurement area first, then the Stimulus are), Repair/Replace the faulty component, and Retest the system. The interviews also uncovered experts' mental model of the test station--in particular, the test station's main functional areas: Measurement Signal, Measurement Data, Stimulus Signal, etc. The expert model is clearly reflected in Sherlock's simulation and coaching.

BEST COPY AVAILABLE

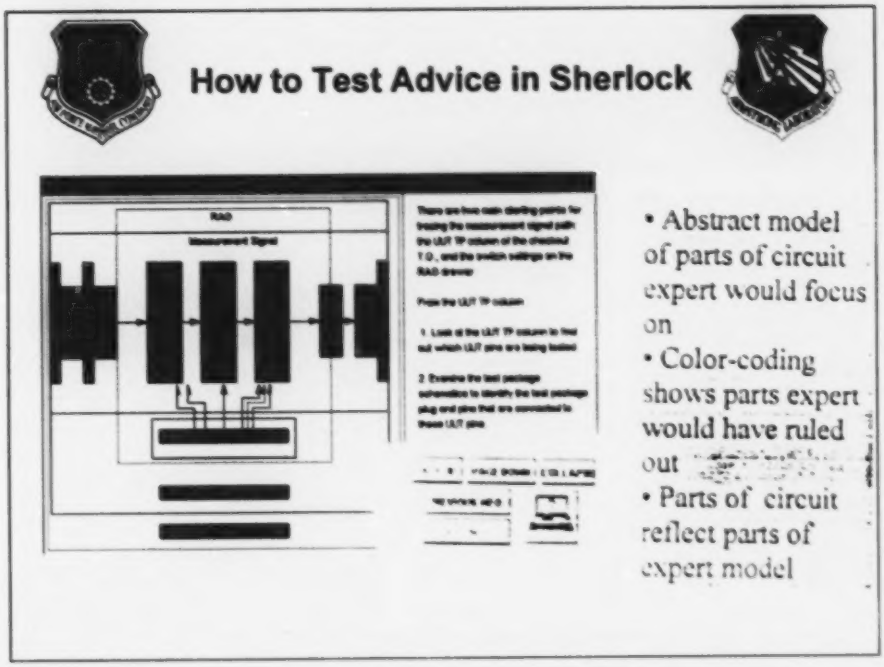


This is an alternative view of the expert model in Sherlock, showing how the model can be instantiated during a given problem scenario. The decomposition into functional areas (Stimulus and Measurement) and components within these areas guided simulation of the test station; this analysis told us what types of components needed to be modeled --e.g., relay cards, logic cards, switches -- as well as which particular components fall into each category. The expert model also drives coaching on what component to investigate next.



The PARI data analysis showed us that three types of knowledge underlie expertise in the avionics job specialty that Sherlock was designed to teach: conceptual (How it works) knowledge, strategic (how to decide what to do and when) knowledge, and procedural (how to do it) knowledge. This classification is reflected in Sherlock's Coaching menu. Students can ask for advice about the circuit as a whole, or about a particular component. They can get information about a particular functional area through the circuit-level how it works option.

BEST COPY AVAILABLE


How to Test Advice in Sherlock
• Abstract model of parts of circuit expert would focus on
• Color-coding shows parts expert would have ruled out
• Parts of circuit reflect parts of expert model

As noted earlier, the expert model is directly reflected in Sherlock's advice -- especially advice about what to test next and why. The color-coded diagram shows students which components and functional areas the expert would rule out at this point. Green means good, red means bad, black means unknown status. Students can select the grey boxes to receive an explanation about the component's status as indicated by its color. The diagrams are abstract representations of the much more complex schematics technicians use on the job and while doing Sherlock problems. These abstractions are meant to portray the expert's mental model of the relevant circuitry. In effect, they give students the message that they need to look past the details and think in terms of broader functional relationships between the components in a circuit. The feedback we received from trainees during field trials of Sherlock suggests that these abstract diagrams are a helpful learning aid.

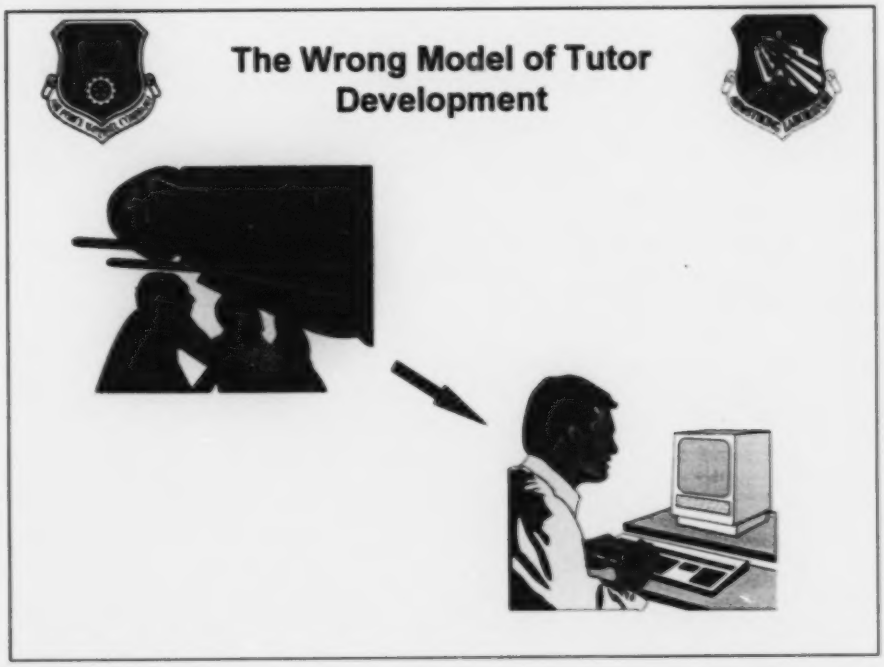
BEST COPY AVAILABLE

Lessons Learned

- **Use database tools from the beginning to structure CTA data into the format the system "understands".**
- **The tutor development model is cyclic, not linear. Input of subject matter experts is critical throughout.**

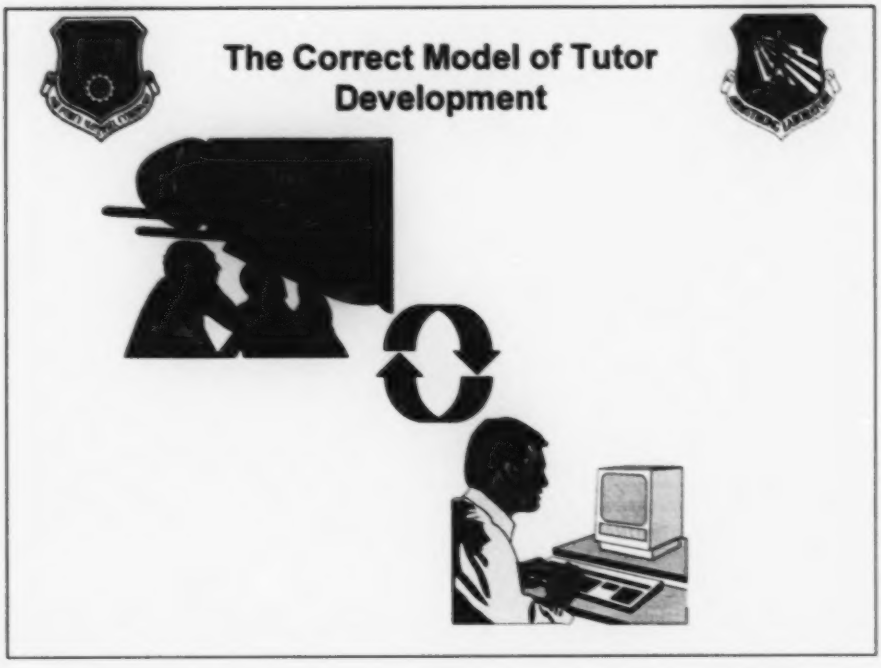
We learned several lessons about using cognitive task analysis to guide tutor development. These are the two we consider most important. We learned the hard way that system developers should work closely with the psychologists conducting the cognitive task analysis early on, in order to devise a way of structuring the data. This structure can then be implemented within a standard database program that task analysts could feasibly use in the field. Doing this would have saved the Sherlock developers a lot of time and aggravation. For example, our programmers puzzled through raw, unstructured transcripts of PARI interviews and lists of hints that they couldn't readily associate with system components. Using a database would have ensured that the data was represented and stored correctly and associated with the right objects. The second point targets what we have noticed to be a common misconception about using CTA for tutor development: that work with subject matter experts ends when the CTA is "done".

BEST COPY AVAILABLE

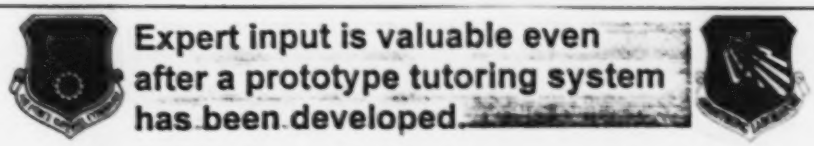


The wrong model of tutor development is that the results of the CTA can be poured directly into the tutoring system and input from subject matter experts is no longer needed.

BEST COPY AVAILABLE



We have learned that the results of the CTA need to be extended and refined throughout the tutor development process. We strongly feel that continuous interaction with subject matter experts was central to the success of the tutoring systems developed under the Basic Job Skills Program.



Experienced Technician's Explanations Gave us Insight into How to Improve System Coaching in Future Tutors. An Example:

Mentor: Allright, on aircraft, a lot of the times, in order to keep something turned off, they will apply the same voltage to both sides. Therefore, there is no current flow. That's how it works.... I was going to let you figure out how these relays actually worked.

Student: Yeah, Well, I guess we learned how they worked today. [Everyone laughs.]

Mentor: I had to prompt you to check all the control voltages out of the A10 card. I shouldn't have done that. That would have really blown your mind. You would've said "gee, you're 28 volts all over the place, why is that?" There's a reason for that. Remember on these cards, there's only one relay selected at a time...If you would have checked the other control lines, you would have found 28 volts everywhere...

In fact, even after Sherlock was deployed, we used the prototype tutor to learn how we could build better systems in the future. To do this, we observed experienced avionics technicians from the Air National Guard "mentoring" students from local avionics technical schools. Students collaborated with a peer on Sherlock problems. They asked their mentor questions, when they could not figure out what to do on their own. After students solved the problem, the mentor debriefed students. One of the many things we learned from these observations is that when experts explain, they don't separate information about how components work from advice about what to do next. Sherlock's advice options make this separation. Instead, the expert technicians consistently integrated system with strategic knowledge, as shown in this example.

This explanation occurred during post-problem debrief. We found that students tended to seek explanations after they solved the problem; during problem solving, students mainly ask what to do next. The mentor was justifying his advice that students should test every data control signal to a relay card. His justification is grounded in knowledge about how the relay card works and how this relates, more generally, to knowledge about how aircraft systems work. In effect, this explanation models an expert's ability to activate the appropriate system knowledge for a given action. We found that explanations like this enabled students to carry out appropriate actions in similar contexts during future problem-solving sessions and to give richer explanations to their peers about why to carry out certain actions.

Mentor: Then you really would have been confused. "NOW what going on?" But that's how they turn it off, they put 28 volts on both sides. Yeah, that keeps it energized, that holds it off. It also makes it reset, too. If you have 28 volts, uhh, on all your control lines and you cut off your source momentarily, you'll do it, it'll give you a reset for the board. That's how it does it. That's how it resets the board. That's what that little diode is in there for...So it, so it's, it's not a stupid machine. It's a lot more sophisticated than you thought it was. Because this way, there is no mess up on selection; on deactivating your relays.

Summary

Cognitive task analysis provides a framework for focusing instruction on critical cognitive skills associated with expertise in complex problem solving tasks, thereby enhancing instructional effectiveness of intelligent tutoring systems

Efficiencies in tutor development can be achieved with tools that allow the direct input of certain data structures from CTA data or by subject matter experts

CTA is an iterative process; tutor development requires continuous involvement of subject-matter experts

BEST COPY AVAILABLE